Künstliche Intelligenz

Einführung, Einsatz-Szenarien sowie Gefahren und Gegenmaßnahmen

Tim Hoffmann

IT4B Digital Summit | 02.07.2025

Datenschutz

Informationssicherheit

Tim Hoffmann

UIMC

Berater für Datenschutz und Informationssicherheit

- Wirtschaftswissenschaften an der Universität-GH Essen
- Studien-Schwerpunkte; u. a.
  - » Organisation
  - » Informationsmanagement
- Seit 2002 als Berater mit den Schwerpunkten
  - » Datenschutz und
  - » Informationssicherheit / IT-Sicherheit

Wer wir sind

UIMC

Datenschutz

Externe Datenschutzbeauftragung, Audits, E-Learning, Dienstleister-Auditierung etc.

Informationssicherheit

Aufbau eines ISMS bis zur Zertifizierung, Informationssicherheitsbeauftragter etc.

Organisation / Strategie

Beratungsleistungen bei konzeptionellen und strategischen Fragestellungen

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

K

KI oder Marketing-Label?
Die Grenzen sind fließend.

*„KI-System“ ein maschinengestütztes System, das für einen in unterschiedlichem Grade **autonomen Betrieb** ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können*
(KI-Verordnung)

- Forschungsgebiet der Informatik: (noch) keine abschließende Definition
- KI kann aus Daten lernen und ihre Leistung verbessern
- bisherige algorithmenbasierte Systeme basieren auf festgelegten Regeln
- Unterscheidung:
  - » Generative KI: erzeugt originelle Daten (Text, Bilder oder Musik)
  - » Diskriminative KI: erkennt Unterschiede zwischen verschiedenen Daten

Szenarien für diskriminative KI

- Qualitätskontrollen (Auswertung großer Datenmengen, Anomalien feststellen)
- Bewertung/Selektion von Kunden & Bewerbern
- Analyse von Programm-Code
- Lieferketten (z. B. Lageroptimierung)
- Unterstützung von Security-Maßnahmen
  - » Erkennung von Cyberangriffen,
  - » Betrugsversuchen (z. B. bei Banken) und von
  - » unerwünschten Inhalten (Spam/Phishing)

Szenarien für generative KI

- Erzeugung von Pressemitteilungen, Werbemitteln, Präsentationen, Protokollen
- Individuelle Bilder statt Stock-Fotos
- Übersetzungen
- Erzeugung von Programm-Code
- Chatbots
- Kollaborative Multiagentensysteme

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Ethische Themen

- Kontrolle und Sicherheit
- Arbeitsplatzverluste
- Verantwortlichkeit und Haftung
- Privatsphäre und Überwachung
- Bias und Fairness
- Existenzielle Risiken
- Menschliche Würde
- Verteilungsgerechtigkeit
- Ethische Entscheidungsfindung
- Menschenrechte

KI-Verordnung

- *Der AI Act (KI-Verordnung) soll die Einführung von **menschen-zentrierten** und **vertrauenswürdigen** KI-Systemen fördern und gleichzeitig ein **hohes Maß an Schutz** für Gesundheit, Sicherheit und Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Schutz der Umwelt gewährleisten.*
- Risikobasierter Ansatz
- Ersetzt nicht die DSGVO und ist nicht DSGVO 2.0

Akteure in KI-Verordnung

- Anbieter („Provider“)
- Produkthersteller („Product Manufacturer“)
- Bevollmächtigter („Authorised Representative“)
- Einführer („Importer“)
- Händler („Distributor“)
- Betreiber („Deployer“)

Akteure in KI-Verordnung

- **Anbieter („Provider“)**
- Produkthersteller („Product Manufacturer“)
- Bevollmächtigter („Authorised Representative“)
- Einführer („Importer“)
- Händler („Distributor“)
- **Betreiber („Deployer“)**

„Hersteller“

Vorsicht bei Customizing:
Ab wann wird Betreiber zum Anwender?

„Nutzer“

Zeitplan der Einführung

Inkrafttreten
01.08.2024

Verbotene KI-Praktiken
KI-Kompetenz
02.02.2025

General Purpose AI
Notifizierungsstellen
Governance
Strafen!!
02.08.2025

Hochrisiko-KI (Anhang III)
KI-Systeme mit geringen und minimalen Risiko
KI-Reallabore einsatzbereit
Sonstige Verpflichtungen
02.08.2026

Hochrisiko-KI (Anhang I)
02.08.2027

Risikostufen von KI-Systemen

inakzeptabel	**Verboten**, weil sie im Widerspruch zu den Werten der EU stehen	Social Scoring, unterschwellige Technologien zur (negativen) Beeinflussung, Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz
Hoch	Anforderungen an Anbieter und Betreiber	Personalmanagement, Kreditwürdigkeit, Vertragsabschlüsse
begrenzt	Risikomanagement durch Transparenzpflichten	Chatbots / generative KI für Bilder, Videos, Stimmen etc.
minimal	Freiwillige Verhaltens-Kodize	Spamfilter, Übersetzung, Rechtschreib- und Grammatikkorrekturen

Verpflichtungen für *Betreiber*

UIMC

Beachte: keine vollständige Auflistung

	Hoch	begrenzt	minimal
KI-Kompetenz	✓	✓	✓
Transparenz gegenüber nachgelagerten Akteuren	✓	✓	
Verwendung des KI-Systems laut Betriebsanleitung	✓		
Menschliche Aufsicht	✓		
Überwachung des KI-Systems	✓		
Meldung von schwerwiegenden Vorfällen	✓		
Aufbewahrung von erzeugten Protokollen	✓		
Sofern relevant: Datenschutz-Folgenabschätzung	✓		
Zusammenarbeit mit zuständigen nationalen Behörden	✓		
Recht auf Erläuterung der Entscheidungsfindung im Einzelfall	✓		
Informationspflichten gegenüber der Arbeitnehmerinnen-Vertretung *sofern Arbeitgeberin Hochrisiko-KI-Systeme am Arbeitsplatz einsetzt*	✓		

Verpflichtungen für *Anbieter*

Beachte: keine vollständige Auflistung

	Hoch	begrenzt	minimal
KI-Kompetenz	✓	✓	✓
Transparenz gegenüber nachgelagerten Akteuren	✓	✓	
Anforderungen an Daten	✓		
Technisch Dokumentation	✓		
Zusammenarbeit mit Behörden	✓		
Risikomanagement	✓		
Genauigkeit, Robustheit und Cybersícherheit	✓		
Registrierungs- und Mitteilungspflichten	✓		
Meldepflichten gegenüber Behörden	✓		
Implementierung von menschlichen Überwachungstools	✓		
Kennzeichnungspflichten, Barrierefreiheitsanforderungen, Aufbewahrungspflichten, Qualitätsmanagement, Korrekturmaßnahmen, Aufzeichnung von Ereignissen, Komformitätserklärung/-kennzeichnung	✓		

Verpflichtungen für *Anbieter*

Beachte: keine vollständige Auflistung	Hoch	GPAI systemisches Risiko	GPAI	begrenzt	minimal
KI-Kompetenz	✓	✓	✓	✓	✓
Transparenz gegenüber nachgelagerten Akteuren	✓	✓	✓	✓	
Anforderungen an Daten	✓	✓	✓		
Technisch Dokumentation	✓	✓	✓		
Zusammenarbeit mit Behörden	✓	✓	✓		
Risikomanagement	✓	✓			
Genauigkeit, Robustheit und Cybersícherheit	✓	✓			
Registrierungs- und Mitteilungspflichten	✓				
Meldepflichten gegenüber Behörden	✓				
Implementierung von menschlichen Überwachungstools	✓				
Kennzeichnungspflichten, Barrierefreiheitsanforderungen, Aufbewahrungspflichten, Qualitätsmanagement, Korrekturmaßnahmen, Aufzeichnung von Ereignissen, Konformitätserklärung/-kennzeichnung	✓				

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Gefahren mittels KI
(eigene Nutzung)

- Bias, Halluzinationen oder andere Integritätsprobleme durch
  - » veraltete oder „vergiftete“ Datenbasis
  - » schlechte Modell-Auswahl
  - » Schwache Daten/fehlende Qualität
- Problematische Nachvollziehbarkeit der Ergebnisse
- Anker-Effekt bei der Nutzung der Ergebnisse

Gefahren mittels KI

IMC

- E
- P

ChatGPT Deutsch

Hier kannst du das Modell GPT-4o mini von OpenAI kostenlos und ohne Registrierung nutzen. Der Chat wird über die offizielle API von OpenAI betrieben, jedoch handelt es sich hierbei nicht um das offizielle Chat-Interface.

Zum DALL-E 3 Bildgenerator wechseln

Hallo Mensch, ich bin ein KI-Chatbot, Modell GPT-4o mini.

Wie viele Kinder hat die ehemalige Bundeskanzlerin Angela Merkel?

Angela Merkel hat keine leiblichen Kinder. Sie ist jedoch Stiefmutter von zwei Kindern aus der ersten Ehe ihres Mannes, Joachim Sauer. Wenn du weitere Fragen hast, stehe ich dir gerne zur Verfügung!

Schreibe hier deine Nachricht...

Powered by ChatGPT API

Gefahren mittels KI
(eigene Nutzung)

- Bias, Halluzinationen oder andere Integritätsprobleme durch
  - » veraltete oder „vergiftete“ Datenbasis
  - » schlechte Modell-Auswahl
  - » Schwache Daten/fehlende Qualität
- Problematische Nachvollziehbarkeit der Ergebnisse
- Anker-Effekt bei der Nutzung der Ergebnisse

Gefahren mittels KI
(eigene Nutzung)

- Schatten-IT durch KI aktuell verstärkt
- Rechtliche Probleme [auch beim Training der KI]
  - » Urheberrechte
  - » Datenschutz
  - » Geschäftsgeheimnisse (eigene / fremde)
  - » Sonstige: Markenrecht, UWG, AGG etc.
- Wichtige Frage:
  - » Werden die eingegebenen Daten auch für das Training der KI genutzt?


Handelsblatt
Anmelden
Abo
Schatten-KI
Sieben von zehn Arbeitnehmern nutzen KI-Werkzeuge ohne Freigabe ihrer Firma
Der Druck bei der Arbeit ist hoc
kann Abhilfe schaffen, aber viel
Einsatz. Beschäftigte handeln e
Stephan Scheuer
09.05.2024 - 15:07 Uhr
San Francisco. Während viele Unternehmen beim Einsatz von Künstlicher Intelligenz (KI) zögern, schaffen Beschäftigte oft Tatsachen. Sieben von zehn Mitarbeitenden (71 Prozent) von Firmen in Deutschland nutzen KI-Werkzeuge bei der Arbeit, ohne dass diese von den Unternehmen bereitgestellt werden.
Das ist das Ergebnis einer Umfrage von Microsoft und dem Karrierenetzwerk LinkedIn unter 31.000 Vollzeitbeschäftigten weltweit. In Frankreich ist der Anteil mit 78 Prozent sogar noch höher. In den USA ist er mit 63 Prozent vergleichsweise gering.
Der Grund für das Vorgehen der Beschäftigten liegt laut der Umfrage zu einem großen Teil in der hohen Arbeitsbelastung. 68 Prozent der Befragten gaben an,
Geschwindigkeit und dem Umfang der Arbeit hadern. 46
Prozent der Befragten
ausgebrannt.
els KI
Nutzung)
UiMC
https://www.handelsblatt.com/technik/ki/schatten-ki-sieben-von-zehn-arbeitnehmern-nutzen-ki-werkzeuge-ohne-freigabe-ihrer-firma/100037174.html

Gefahren mittels KI
(eigene Nutzung)

- Schatten-IT durch KI aktuell verstärkt
- Rechtliche Probleme [auch beim Training der KI]
  - » Urheberrechte
  - » Datenschutz
  - » Geschä
  - » Sonstig
- Wichtige F
  - » Werden
    das Tra

Urheberrechte:

- ✓ Werk: KI-Output in der Regel nicht geschützt
- ✓ Daten für das Training dürfen nicht urheberrechtlich geschützt sein
- ✓ Wenn Output zu nah am Original: Urheberrechtsverletzung

Gefahren mittels KI
(eigene Nutzung)

- Schatten-IT durch KI aktuell verstärkt
- Rechtliche Proble
  - » Urheberrechte
  - » Datenschutz
  - » Geschäftsgehei
  - » Sonstige: Marke
- Wichtige Frage:
  - » Werden die eing
    das Training der

Datenschutz:

- ✓ Rechtmäßigkeit der Datenverarbeitung (u. a. Rechtsgrundlage)
- ✓ Zweckbindung
- ✓ Datenminimierung
- ✓ Richtigkeit der Daten
- ✓ Speicherbegrenzung
- ✓ Vertraulichkeit/Integrität der Datenverarbeitung
- ✓ Betroffenenrechte (Info-Pflichten, Auskunft, Löschung etc.)
- ✓ Weitergabe der Daten (AVV, JC, Übermittlung)
- ✓ Datenschutz-Folgenabschätzung?!

Gefahren mittels KI
(eigene Nutzung)

- Manipulation der eingesetzten KI
  - » Evasion Attack (Manipulation der Eingabe)
  - » Poisening Attacks (Vergiftung der Trainingsdaten)

Gefahren mittels KI
(Nutzung durch Dritte/Fremde)

- Unterstützung bei Schwachstellensuche
- Schnellere Analyse von Patches
  (Erhöhung von „Zero Days“?)
- Re-Identifizierung von anonymisierten Daten
- Generierung / Verbesserung / Individualisierung von Malware

Gefahren mittels KI
(Nutzung durch Dritte/Fremde)

- Wissenssammlung und -aufbereitung im Rahmen von Cyberangriffen
- Fake-News / Hoaxes
- Professionalisierung von Social Engineering

Exkurs „Social Engineering“

„soziale Manipulation“

- ist an sich nichts Neues (früher „Trickbetrug“)
- Cyber-Kriminelle verleiten das Opfer auf diese Weise beispielsweise dazu,
  - » vertrauliche Informationen preiszugeben,
  - » Sicherheitsfunktionen auszuhebeln,
  - » Überweisungen zu tätigen oder
  - » Schadsoftware auf dem Gerät zu installieren.

Exkurs „Social Engineering“

- Ausnutzung menschlicher Eigenschaften, wie z. B. Hilfsbereitschaft, Vertrauen, Angst oder Respekt vor Autorität.
- Möglichkeiten / Varianten:
  - » Soziale Netzwerke
  - » Malvertising
  - » USB-Drop
  - » Dumpster Diving
  - » Phishing
  - » Tailgaiting

Achtung, Fake News: Donald Trump in New York verhaftet

Donald Trump läuft vor Polizeibeamten weg und Wladimir Putin kniet vor Xi Jinping? Die Deep Fakes sind täuschend echt. Fast jedenfalls.

Franz Becchi

24.03.2023 | 11:12 Uhr

Der frühere US-Präsident Donald Trump wird in New York verhaftet: Das Bild ist nicht echt. Hergestellt von KI Midjourney

https://www.berliner-zeitung.de/politik-gesellschaft/virale-bilder-deepfake-fotos-zeigen-wie-donald-trump-verhaftet-wird-li.330828

Freddie Mercury AI - I Will Always Love You

https://www.youtube.com/watch?v=zDaTFLweCcs

Angebliche Kapitulation

Meta löscht Fakevideo, das Selenskyj falsche Worte in den Mund legt

Im Netz kursiert ein Deepfake-Video, in dem Präsident Selenskyj die Ukrainer zur Kapitulation aufruft. Der Clip wurde schnell als Fälschung enttarnt – auch durch ein Statement des echten Selenskyj.

17.03.2022, 09.48 Uhr

https://www.youtube.com/watch?v=X17yrEV5sl4

Wie ein cleverer Ferrari-Manager einen Deepfake entlarvte

30. Juli, 2024 09:20

https://www.it-daily.net/shortnews/wie-ein-cleverer-ferrari-manager-einen-deepfake-entlarvte

C+ KI-BETRUGSMASCHEN

Der falsche Bill: So wurde der Bayer-Chef Opfer eines Deepfakes

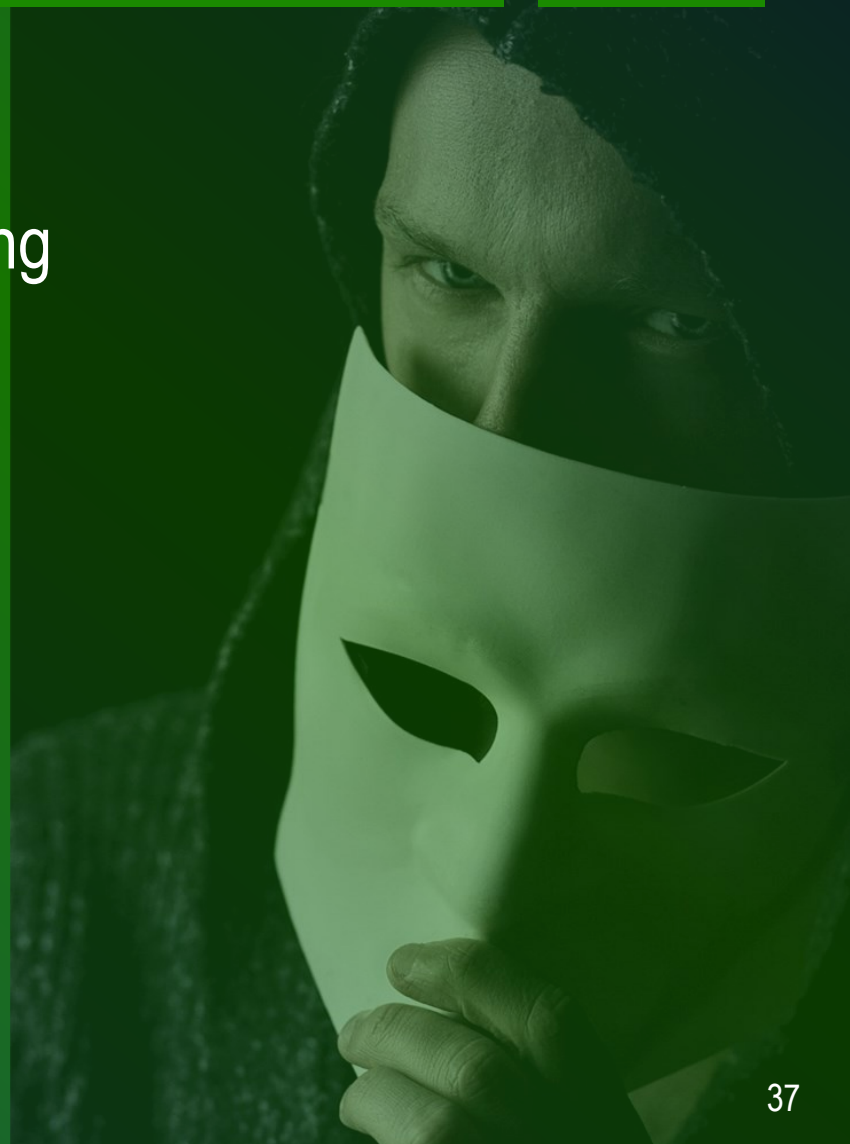
Der echte Bill Anderson: Der Bayer-CEO wurde Opfer ein

© Henning Kaiser/dpa / Picture Alliance

https://www.capital.de/wirtschaft-politik/bayer-chef-wird-deepfake-opfer--ki-betrug-haeuft-sich-35486154.html

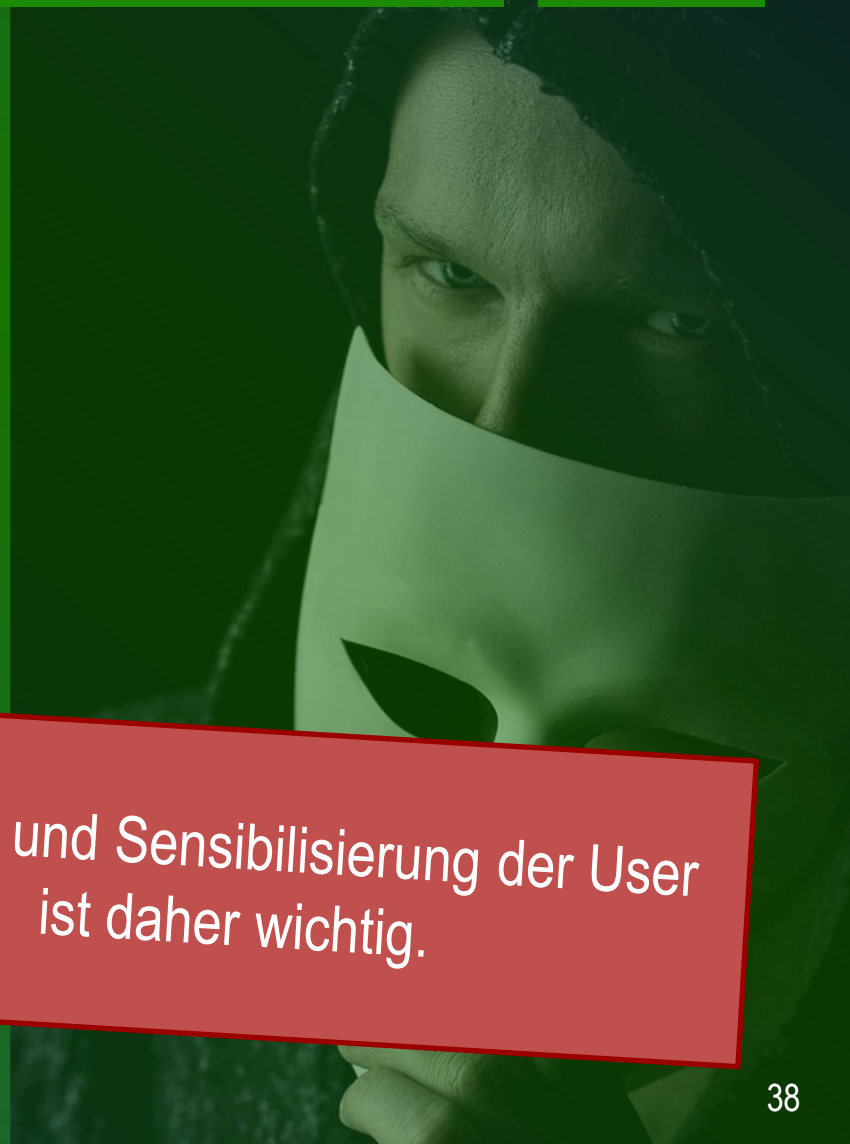
KI-Szenarien im Social Engineering

- Text:
  - » Größere Masse bei gleichzeitiger Individualisierung
- Stimme:
  - » Text-zu Sprache (TTS, Text-to-Speech)
  - » Stimmenkonvertierung (VC, Voice Conversion)
- Fotos/Videos
  - » Face Swapping und Face Reenactment



Potentiale durch KI

- Text: Verbesserung von Phishing-Mails, CEO-Frauds etc.
- Stimme: Fake-Anrufe
- Foto / Video:
  - » Aufbau von Vertrauen
  - » Fake-Botschaften
  - » Eindringen in Videokonferenzen
  - » Überwindung biometrischer Systeme



Schulung und Sensibilisierung der User ist daher wichtig.

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Gegenmaßnahmen

- Drei Dinge
  - » Sensibilisierung der Mitarbeiter:innen
  - » Sensibilisierung des Managements
  - » Sensibilisierung aller

Bedeutsam auch im Hinblick auf richtige Nutzung (Halluzinationen etc.) sowie auf Angriffsszenarien.

„Die Anbieter und Betreiber von KI-Systemen ergreifen **Maßnahmen**, um nach besten Kräften sicherzustellen, **dass ihr** Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, **über ein ausreichendes Maß an KI-Kompetenz verfügen**, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.“ *(Artikel 4 KI-VO)*

Gegenmaßnahmen

- Drei Dinge
  - » Sensibilisierung der Mitarbeiter:innen
  - » Sensibilisierung des Managements
  - » Sensibilisierung aller
- Aber auch
  - » Risikobewertung (anpassen)
  - » Organisation anpassen (z. B. BCM)
  - » aufbauende Prozesse (z. B. Codewords, 4-Augen-Prinzip)

Gegenmaßnahmen

- IT-Sicherheitsmaßnahmen
  - » Patchmanagement
  - » Aufbau einer resilienten IT-Infrastruktur
  - » Verbesserung der Angriffserkennung
  - » Multi-Faktor- Authentifizierung
  - » eigene Nutzung der KI für Verteidigungsmaßnahmen (z. B. Erkennung von Bedrohungen und Schwachstellen).

Gegenmaßnahmen

- Schutz der KI
  - » Sicherstellung der Qualität und der Integrität der Trainingsdaten
  - » Schutz der Modelle vor Diebstahl und Manipulation
  - » Durchführung von umfassenden Tests
  - » Sorgfältige Auswahl des KI-Systems und des betreibenden Unternehmens
  - » etc.

Agenda

KI: Eine kurze Einführung und Einsatz-Szenarien

Rechtliche Anforderungen: KI-Verordnung

Gefahren bei der Nutzung / Bedrohungen durch KI

Gegenmaßnahmen

Fazit

Fazit

- KI betrifft alle Unternehmen, egal ob sie es aktiv nutzen
- KI bietet viele Chancen (insb. im Hinblick auf die Effizienz)
- Vor dem internen Einsatz
  - » Betrachtung rechtlicher Implikationen
  - » Durchführung einer Risikobewertung
  - » Umsetzung angemessener Sicherheitsmaßnahmen
- Risikobewertung und Anpassung des ISMS
- A & O: Belegschaft ausreichend schulen und sensibilisieren

Wie wir helfen können

- Durchführung einer Risikobewertung
  (zur Einschätzung, welche Vorgaben der KI-Verordnung zu beachten sind)
- Erstellung von Richtlinien und weiteren Strukturen
  (zum rechtkonformen und sicheren Einsatz von KI)
- Schaffung von KI-Kompetenz
  (u. a. durch Workshops oder Bereitstellung eines E-Learning-Kurses)
- Überarbeitung Ihrer Informationssicherheits-Organisation
  (Modifikation Ihres ISMS durch die neue Bedrohungslage durch KI)

F.A.Q.

- Wie schulen?
  - » E-Learning, Workshops, Präsenzschulungen
- Nachweis der Schulungen?
  - » Teilnehmerlisten, Tests, technische Protokollierung
- KI-Beauftragter nötig?
  - » Dies ist gemäß KI-Verordnung nicht explizit gefordert
- Wer haftet?
  - » Allgemeine Haftungsregelungen (z. B. § 130 OWiG)

Fragen??

UIMC

Mancher ertrinkt lieber,
als daß er um Hilfe ruft.
– Wilhelm Busch

thoffmann@uimc.de

Folgen Sie mir auf LinkedIn:
www.linkedIn.com/in/tim-hoffmann-uimc

Diese und weitere Präsentationen unter
http://update.uimcollege.de
Gastzugang: web.e#2025*UIMC

Kontaktdaten

UIMC Dr. Vossbein GmbH & Co. KG
Otto-Hausmann-Ring 113
42115 Wuppertal
Telefon: +49 202 946 7726 300
Telefax: +49 202 946 7726 9300
E-Mail: consultants@uimc.de
Internet: www.UIMC.de

UIMCert GmbH
Otto-Hausmann-Ring 113
42115 Wuppertal
Telefon: +49 202 946 7726 300
Telefax: +49 202 946 7726 9300
E-Mail: certification@uimcert.de
Internet: www.UIMCert.de

Wuppertal

Saarbrücken

Berlin

Wien

IT-Sicherheit mit System

Aufbau eines Informationssicherheits-Managementsystem

Schulung der Mitarbeiter

Auditierung der IT-Sicherheit

Risiko-Workshop

u. v. m.

… bis hin zur Zertifizierungsreife

Datenschutz von A bis Z

UIMC®

Externe Datenschutzbeauftragung

Schulung der Mitarbeiter

Datenschutz-Checkup / Auditierung

Datenschutz-Management / -Organisation

u. v. m.

Pragmatischer Datenschutz weiterhin möglich